



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/657,421	09/08/2003	Jordan Cohen	2639/C24	9023
2101 7590 11/19/2010 Sunstein Kann Murphy & Timbers LLP 125 SUMMER STREET BOSTON, MA 02110-1618				
EXAMINER SHAH, PARAS D				
ART UNIT 2626		PAPER NUMBER		
NOTIFICATION DATE 11/19/2010		DELIVERY MODE ELECTRONIC		

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

PATENTS@SUNSTEINLAW.COM



UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

**BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES**

Application Number: 10/657,421
Filing Date: September 08, 2003
Appellant(s): COHEN ET AL.

Eric L. Prah
Wilmer Cutler Pickering Hale and Dorr LLP
60 State Street
Boston, Massachusetts 02109
For Appellant

EXAMINER'S ANSWER

This is in response to the appeal brief filed 09/22/2010 appealing from the Office action mailed 09/30/2009.

(1) Real Party in Interest

The examiner has no comment on the statement, or lack of statement, identifying by name the real party in interest in the brief.

(2) Related Appeals and Interferences

The examiner is not aware of any related appeals, interferences, or judicial proceedings which will directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

(3) Status of Claims

The following is a list of claims that are rejected and pending in the application:

Claims 1-4, 6-9, 11, 12, and 15 are pending and have been rejected.

(4) Status of Amendments After Final

The examiner has no comment on the appellant's statement of the status of amendments after final rejection contained in the brief.

(5) Summary of Claimed Subject Matter

The examiner has no comment on the summary of claimed subject matter contained in the brief.

(6) Grounds of Rejection to be Reviewed on Appeal

The examiner has no comment on the appellant's statement of the grounds of rejection to be reviewed on appeal. Every ground of rejection set forth in the Office action from which the appeal is taken (as modified by any advisory actions) is being maintained by the examiner except for the grounds of rejection (if any) listed under the subheading "WITHDRAWN REJECTIONS." New grounds of rejection (if any) are provided under the subheading "NEW GROUNDS OF REJECTION."

(7) Claims Appendix

The examiner has no comment on the copy of the appealed claims contained in the Appendix to the appellant's brief.

(8) Evidence Relied Upon

5796916	Meredith	8-1998
6081780	Lumelsky	6-2000
EP 1271469	Marasek et al.	01-2003
WO 02/097590A	Cameron	12-2002

(9) Grounds of Rejection

The following ground(s) of rejection are applicable to the appealed claims:

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

Claims 1-3, 6-7, 9, 12, and 15 are rejected under 35 U.S.C. 103(a) as being unpatentable over Marasek *et al.* (EP 1 271 469) in view of Meredith (US 5,796,916, issued on 08/18/1998) in view of Lumelsky (US 6,081,780, issued on 06/27/2000) in view of Cameron (WO 02/097590 A, published on 12/05/2002).

As to claims 1 and 9, Marasek teaches a method and system for speech synthesis comprising:

receiving a spoken utterance (see Abstract Figure 1, S1, receive speech input S1) (e.g. It is obvious that a microphone is used to input speech in the system.);

in response to receiving a spoken utterance (see Figure 1, S1, speech is received and corresponding processing performed as shown in the steps below S1):

extracting one or more prosodic parameters (see Figure 1, S21m, extract prosodic features) (e.g. It is obvious that a signal processor would be used to extract prosody features such as described un [0042] and is well known in the art.) from the spoken utterance;

performing speech recognition on the spoken utterance to generate a recognized word (see Figure 1, steps S12 and [0040] recognize speech)

generating a prosodic mimic word using (Figure 1, step S40 and S50 and [0046], speech synthesis is performed on the input speech by applying prosody to a given text (see [0003]) and the one or more prosodic parameters (see Figure 1, step S21, prosody parameters are extracted and applied to storage personality pattern as seen in Figure 1, step S30).

However, Marasek does not specifically teach the alignment of the spoken utterance and the synthesized word.

Meredith does disclose the alignment of the spoken utterance to the synthesized speech (see Abstract).

It would have been obvious to one of ordinary skilled in the to art at the time the invention was made to have combined the speech synthesis for an utterance as presented by Marasek by the alignment of the utterance and the synthesized word presented by Meredith. The motivation to have combined the two references includes the improvement in intonation (see Meredith col. 3, lines 5-10).

However, Marasek in view of Meredith do not specifically teach the generation of a nominal word.

Lumelsky does teach synthesizing a nominal word (e.g. The applicant refers to the nominal word as synonymous to synthesized word) corresponding to the recognized word (see col. 13, lines 29-41, synthetic speech is produced

based on a pre-stored voice selected by the narrator. Further in col. 16, lines 45-65, the speech that has been output can be reconfigured by editing or changing the prosody parameters.); and

It would have been obvious to one of ordinary skilled in the to art at the time the invention was made to have combined the speech synthesis for an utterance as presented by Marasek in view of Meredith by the generation of a default voice output as taught by Lumelsky. The motivation to have combined the references involves editing or altering of output based on user preference (see Lumelsky, col. 16, lines 42-65).

However, Marasek in view of Meredith in view of Lumelsky do not specifically disclose the system implemented on a handheld device and at least one of a command to be executed by the handheld device and a name to be dialed by a handheld device and if the recognized word includes the command, executing the command on the handheld device, and if the recognized word includes the name, dialing a number corresponding to the name.

Cameron does disclose the speech synthesis implemented on a handheld device (see page 5, 6th paragraph and see page 29, 1st paragraph) (e.g. portable is synonymous to handheld and PDA is a handheld device)

at least one of a command to be executed by the handheld device (see page 18, sect. 6, first paragraph, lines 1, command dial makes voice assistant to dial the telephone) and a name to be dialed by a handheld device (see page 18, sect. 6, 2nd paragraph, lines 5) and;

if the recognized word includes the command, executing the command on the handheld device (see page 18, sect. 6, first paragraph, lines 1, command dial makes voice assistant to dial the telephone), and if the recognized word includes the name, dialing a number corresponding to the name (see page 18, sect. 6, 2nd paragraph, lines 5-9 calls John at business number using autodial).

It would have been obvious to one of ordinary skilled in the to art at the time the invention was made to have combined the speech synthesis for an utterance as presented by Marasek in view of Meredith in view of Lumelsky by the implementation on a handheld device for the purpose of portability, which allows the user to use the device anywhere as is apparent and seen in navigation and translation devices, which incorporate speech recognition and generate a synthetic speech output based on user selection (see Cameron page 5, last paragraph, example of recognition and voice output is described and page 10, bullet 10-page 11, command recognition and speech synthesis) for minimal hand/eye distraction (see Cameron, page 32, 2nd paragraph).

As to claim 9, Marasek in view of Meredith in view of Lumelsky in view of Cameron teaches all of the limitations as in claim 1, above. Furthermore, Cameron teaches the use of a processor in for the recognition of command and dialing a number (see Figure 1, CPU)

As to claim 2, Marasek in view of Meredith in view of Lumelsky in view of Cameron teaches all of the limitations as in claim 1, above.

Furthermore, Marasek teaches wherein the one or more prosodic parameters include pitch (see [0042], pitch).

As to claim 3, Marasek in view of Meredith in view of Lumelsky in view of Cameron teaches all of the limitations as in claim 1, above

Furthermore, Marasek teaches wherein the one or more prosodic parameters include timing (see [0042], speech element duration).

As to claim 4, Marasek in view of Meredith in view of Lumelsky in view of Cameron teaches all of the limitations as in claim 1, above.

Furthermore, Marasek teaches wherein the one or more prosodic parameters include energy (see [0042], loudness).

As to claim 6, Marasek in view of Meredith in view of Lumelsky in view of Cameron teaches all of the limitations as in claim 1, above

Furthermore, Meredith teaches comprising temporally (see col. 4, lines col. 4, lines 37-53) (e.g. The reference indicates the use of intervals and a pitch point marking) aligning phones (see col. 3, line 5) (e.g. Phones are synonymous to phonetic symbols) of the spoken utterance and phones of the nominal word (see Abstract).

As to claim 7, Marasek in view of Meredith in view of Lumelsky in view of Cameron teaches all of the limitations as in claim 1, above

Furthermore, Marasek teaches comprising converting the prosodic mimic word into a corresponding audio signal (see Figure 1, steps S40 and S50 and [0048], synthetic speech is output) (e.g. It is obvious that the signal is in audio form in order for the user to listen to the speech generated).

As to claim 12, Marasek in view of Meredith in view of Lumelsky in view of Cameron teaches all of the limitations as in claim 9, above

Furthermore, Lumelsky teaches a storage device (see col. 17, line 22, dsp) including executable instructions (see col. 17, line 21) for speech analysis and processing (see col. 17, lines 17-20, dsp).

As to claims 8, Marasek in view of Meredith in view of Lumelsky in view of Cameron teaches all of the limitations as in claim 1, above

Furthermore, Cameron teaches the use of a portable telephone (see page 5, paragraph 6, line 4) input device (see page 5, paragraph 6, line 1) and the prosodic mimic word (synthesis and presentation of commands to the user) (see Abstract and page 18, paragraph 2, lines 1-8) is provided to a telephone output device (see page 5, paragraph 6, line 2). Further, Cameron discloses the use of a user interface (see Abstract) utilizing a mobile phone (see page 18, line 7 and

page 5, paragraph 5, line 4) (e.g. It is inherent that a portable telephone encompasses a mobile telephone).

As to claim 15, Marasek in view of Meredith in view of Lumelsky in view of Cameron teaches all of the limitations as in claim 1, above

Furthermore, Cameron teaches wherein the command is any one of a plurality of available commands (see page 18, sect. 6, first paragraph, lines 1, command dial makes voice assistant to dial the telephone and see page 10, sec. 10, lines 1, where another command search is described and see page 11, bullet c, lines 2, temporal command, hence plurality of commands are utilized)

(10) Response to Argument

Claims 1-4, 6-9, 11, 12, and 15 are rejected under 35 U.S.C. §103 under Marasek in view of Meredith in view of Lumelsky in view of Cameron.

Appellant asserts on page 8 with respect to claims 1 and 9:

"We note, however, that the examiner's characterization of the teachings of Marasek is not accurate. Marasek does not perform speech synthesis on input speech, he performs it when generating speech in, for example, a dialogue system (e.g. see [0024]), which involves producing speech from templates or stored responses. Marasek never even hints at performing speech synthesis to reproduce the very same speech (or word) that was received and in response to receiving that speech (or word). That would involve receiving an utterance and then playing that received utterance back as synthesized speech, something that Marasek does not do and is not described by Marasek. "

The Examiner respectfully disagrees with the Appellant. The Examiner notes that the Appellant is arguing limitations which are not being claimed. Specifically, the speech synthesis to reproduce the very same speech or word that was received. The current claim recites in lines 9 and 10, that the synthesized nominal word is created from the recognized word. Furthermore, in lines 11 and 12, the prosodic mimic word is generated from the synthesized nominal word and the extracted prosodic parameter. The claims do not recite that the same speech or word is reproduced but rather the claims recite the synthesized and prosodic words being generated based on the recognized or synthesized nominal word, which is different in scope. It is noted that the features upon which appellant relies (i.e., the same speech or word is reproduced) are not recited in the rejected claim(s). Although the claims are interpreted in light of the specification, limitations from the specification are not read into the claims. See *In re Van Geuns*, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993).

Appellant asserts on page 9 with respect to claims 1 and 9:

"There is no need in Marasek to synthesize the very same speech that was just received since it already most accurately reflects the sounds of the speaker, i.e., it already has the prosody of the speaker. Certainly, Marasek has not provided a reason for adopting such an approach.

It is worth reiterating a major difference between the claimed invention and the teachings of Marasek. Marasek says nothing about performing a number of steps in response to receiving a spoken utterance, including generating a recognized word from that spoken utterance and then synthesizing that recognized word to generate a synthesized word. More specifically, Marasek says nothing about performing at least two steps "in response to receiving the spoken utterance," wherein those steps involve: "performing speech recognition on the spoken utterance to generate a recognized word" and then "from the recognized word that is generated from the speech recognition, synthesizing a nominal word." Marasek simply says that the extracted prosody can be

used later to synthesize words which have the personality of a particular speaker."

The Examiner respectfully disagrees with the Appellant. Marasek teaches and suggests that speech received by the system is recognized (see Figure 1, s12 and prosody parameters are extracted (see Figure 1, s21). This is then used to synthesize information in order to produce the speech output (see Figure 1, s40 and s50). In the cited portion presented by the Appellant (see Marasek, [0006]), although Marasek indicates the usage of the personality pattern later on, Marasek further discloses the parameters extracted can be used to *reconstruct* a speech output that *mimics the speech input* and its speaker. Furthermore, the Appellant's indicate that the system uses the pattern to later on perform synthesis. However, with respect to the Appellant's claims, it cannot be seen how the claimed limitations are performed at the same time and not later on. This can be seen by the series of processing steps that take place, for example, the mentioned limitations from lines 5-10. Thus, in order for a prosodic mimic word to be generated, the processing steps must take place, which makes the synthesis to also occur later on. Furthermore, as noted in the prior paragraph, the claims do not recite the synthesis of the very same speech.

Furthermore, the Appellants have not addressed the reference of Lumelsky for which the synthesis of the nominal word has been relied upon. Lumelsky discloses in col. 13, lines 23-28, where the user has selection capability a voice to listen to the synthesized speech. In other words, a default synthesis of text based on user selection. In col., 16, lines 55-65, Lumelsky further discloses the editing or modification of prosody parameters based on user satisfaction of the playback. Hence, Lumelsky discloses the

Art Unit: 2626

synthesis of text to a default form and further modifying the default selection by changing the prosody parameter, where the primary reference of Marasek teaches the generation of a prosodic mimic word from the recognized input speech and extracted prosody parameters.

(11) Related Proceeding(s) Appendix

No decision rendered by a court or the Board is identified by the examiner in the Related Appeals and Interferences section of this examiner's answer.

For the above reasons, it is believed that the rejections should be sustained.

Respectfully submitted,

/P. S./

Examiner, Art Unit 2626

11/11/2010

Conferees:

Paras Shah

/Paras Shah/

Examiner, Art Unit 2626

David Hudspeth

/David R Hudspeth/

Art Unit: 2626

Supervisory Patent Examiner, Art Unit 2626

James Wozniak

/James S. Wozniak/

Supervisory Patent Examiner, Art Unit 2626